মনোবৈজ্ঞানিক অভীক্ষা ও শিক্ষাগত অভীক্ষার মূল পার্থক্য হল উভয়ের উদ্দেশ্য পৃথক। শিক্ষাগত অভীক্ষার সাহায্যে ব্যক্তির শিক্ষাগত পারদর্শীতা, পাঠ্যবিষয়ক দুর্বলতা, পঠন ক্ষমতা ইত্যাদি পরিমাপ করা হয় কিন্তু মনোবৈজ্ঞানিক অভীক্ষার দ্বারা ব্যক্তির মানসিক বৈশিষ্ট্য যেমন—বুদ্ধি, আগ্রহ, প্রবণতা, সৃজনশীলতা ইত্যাদি পরিমাপ করা হয়।

❂ ২.৩ নির্ভরযোগ্যতা: অর্থ, প্রকারভেদ ও পরিমাপ (Reliabity: Meaning, types and Measurement) :

❑ ২.৩.১ নির্ভরযোগ্যতা (Reliability) :

❖ ২.৩.১.১ সংজ্ঞা ও ধারণা (Definition and concept) :

ধরা যাক, কোনো অভীক্ষার অন্তর্গত অভীক্ষাপদগুলির বেশিরভাগই কোনো শিক্ষার্থী অনুমানের ভিত্তিতে উত্তর দিয়েছে এবং সেই অনুমান সঠিক হওয়ায় ভালো স্কোর পেয়েছে। এক্ষেত্রে শিক্ষার্থীর প্রাপ্ত স্কোর তার প্রকৃত সম্পাদনক্ষমতা বা পারদর্শিতার আস্থাযোগ্য পরিমাপ নয়। আর একটি ক্ষেত্রে ধরা যাক, ঐ শিক্ষার্থী অভীক্ষার অন্তর্গত অভীক্ষাপদগুলির উত্তর জানা থাকলেও প্রশ্নের জটিল ভাষা ও দ্ব্যর্থক বিবৃতির জন্য অনিচ্ছাকৃতভাবে ভুল উত্তরদানে চালিত হয়েছে এবং নিম্নমানের স্কোর অর্জন করেছে। কিংবা এমনও হতে পারে, ঐ শিক্ষার্থী সমগ্র পাঠ্যসূচির নির্বাচিত কিছু অংশ (Selective study) পাঠ করেছে এবং ভাগ্যবশত অভীক্ষার বেশিরভাগ প্রশ্ন সেখান থেকেই আসায় অপেক্ষাকৃত ভালো স্কোর অর্জন করার মত উত্তর দিয়েছে। উপরের তিনটি ক্ষেত্রে ঐ শিক্ষার্থীর প্রাপ্ত স্কোরের পরিবর্তনশীলতার জন্য তার পারদর্শিতা সম্পর্কে নির্ভরযোগ্য পরিমাপ পাওয়া সম্ভব নয়।

যদি কোনো শিক্ষাগত বা মনোবৈজ্ঞানিক অভীক্ষা দ্বারা একই শিক্ষার্থীর একই বৈশিষ্ট্য প্রায় একই পরিস্থিতিতে একাধিকবার পরিমাপ করলে একই ফল সবসময় পাওয়া যায় তাহলে অভীক্ষাটিকে এবং প্রাপ্ত পরিমাপটিকে নির্ভরযোগ্য (Reliability) বলা যায়। পরিমাপের এই সামঞ্জস্যপূর্ণ স্থিরতা বা অপরিবর্তনীয়তার মাত্রাকেই পরিমাপটির **নির্ভরযোগ্যতা (Reliability) বলা হয়।**

শিক্ষামনোবিদ্গণের মতে— Reliability of a test may be defined as the degree to which a test is consistent, stable, dependable or trustworthy in measuring what it is measuring."

J.C. Aggarwal নির্ভরযোগ্যতার ধারণা দিতে গিয়ে বলেছেন– "By this term we mean the degree of accuracy with which an examination measures what it seeks to measure."

Scanned with CamScanner

শিক্ষাবিদ্ Frank. S. Freeman এর মতে— "The reliability of test is its stability to yield consistent result from one set of measures to another."

❖ **২.৩.১.২ নির্ভরযোগ্যতা নির্ণয়ের পদ্ধতি (Methods of determining Reliability) :**

কোনো শিক্ষাগত বা মনোবৈজ্ঞানিক অভীক্ষার নির্ভরযোগ্যতা নির্ণয় করার জন্য দু'ধরনের রাশিবৈজ্ঞানিক কৌশল ব্যবহার করা হয়। যথা—

(i) সহগতির সহগাঙ্ক (Coefficient of Correlation) নির্ণয়।
(ii) আদর্শ ত্রুটি (Standard Error) নির্ণয়।

সহগতির সহগাঙ্ক নির্ণয়ের মাধ্যমে কোনো অভীক্ষার নির্ভরযোগ্যতা সংরক্ষণের চারটি উল্লেখযোগ্য পদ্ধতি হল—

(A) পুনঃঅভীক্ষা পদ্ধতি (Test-Retest Method)
(B) বিকল্প পদ্ধতি বা সমান্তরাল পদ্ধতি বা সমতুল্য অভীক্ষা পদ্ধতি (Alternative Form or Parallel Form or Equivalent Test Method)
(C) অর্ধ-দ্বিখণ্ডিত পদ্ধতি (Split-half Method)
(D) অভ্যন্তরীণ-অপরিবর্তনীয় পদ্ধতি (Internal Consistency Method)

উপরোক্ত পদ্ধতিগুলি সম্পর্কে নিম্নে আলোচনা করা হল—

২.৩.১.২.১ পুনঃঅভীক্ষা পদ্ধতি (Test-Retest Method) :

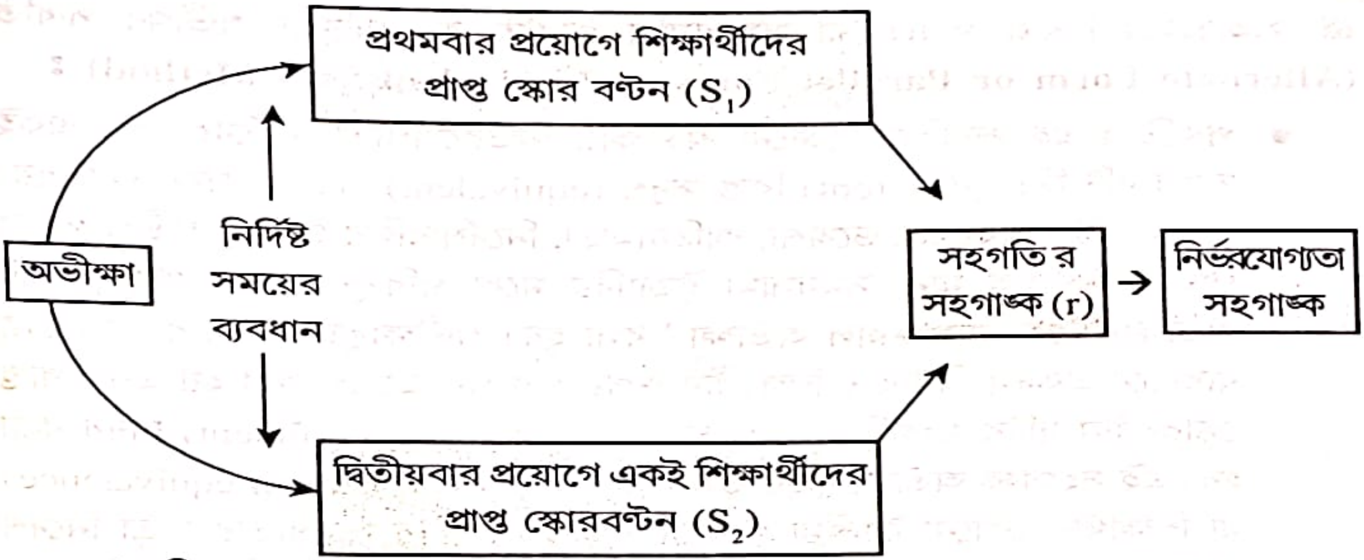
- **পদ্ধতি :**

  কোনো অভীক্ষার নির্ভরযোগ্যতা নির্ণয়ের সবচেয়ে সহজ পদ্ধতি হল পুনঃঅভীক্ষা পদ্ধতি। এই পদ্ধতিতে বিবেচ্য অভীক্ষাটি (যার নির্ভরযোগ্যতা নির্ণয় করতে হবে) একদল শিক্ষার্থীর উপর প্রায় একই পরিস্থিতিতে শুধুমাত্র সময়ের ব্যবধান পরপর দু'বার প্রয়োগ করা হয়।

  দুটি ভিন্ন পরিস্থিতিতে শিক্ষার্থীদের প্রাপ্ত স্কোর বণ্টন দুটির মধ্যে সহগতির সহগাঙ্ক (Correlation Coefficient) নির্ণয় করা হয়।

  এই সহগতির সহগাঙ্কই হল অভীক্ষাটির নির্ভরযোগ্যতার সহগাঙ্ক (Reliability Coefficient)। এই নির্ভরযোগ্যতা সহগাঙ্ক প্রাপ্ত স্কোরের স্থায়িত্ব (Stability) নির্দেশ করে।

Scanned with CamScanner



- **ত্রুটি ঃ**

পুনঃঅভীক্ষা পদ্ধতিতে নির্ভরযোগ্যতা নির্ণয়ে নিম্নলিখিত ত্রুটিগুলি লক্ষ করা যায়। যথা—

(i) অভীক্ষাটির দুবার প্রয়োগের মধ্যে যদি সময়ের ব্যবধান খুব কম হয় (1-2 দিন), তাহলে অভীক্ষার্থীর স্মৃতি, অনুশীলন, অভিজ্ঞতা সঞ্চালনের জন্য দ্বিতীয় প্রয়োগে প্রাপ্ত ফলাফলে উন্নতি লক্ষ করা যায়। ফলে নির্ভরযোগ্যতা সহগাঙ্কের যে মান পাওয়া যায় তা অভীক্ষাটির নির্ভরযোগ্যতার সঠিক পরিমাপ নয়।

(ii) আবার সময়ের ব্যবধান খুব বেশি হলেও (6 মাস বা তার বেশি) শিক্ষার্থীর অর্জিত নতুন অভিজ্ঞতাগুলি দ্বিতীয় প্রয়োগের ফলাফলকে প্রভাবিত করবে। ফলে নির্ভরযোগ্যতা সহগাঙ্কের মান হ্রাস পাবে।

(iii) অভীক্ষা প্রয়োগের ক্ষেত্রে সমপরিস্থিতি বজায় রাখা অনেকক্ষেত্রে সম্ভব না হওয়ায় নির্ভরযোগ্যতার সঠিক মান পাওয়া যায় না।

অর্থাৎ উপরোক্ত ত্রুটিগুলির জন্য পুনঃঅভীক্ষা পদ্ধতিতে কোনো অভীক্ষার নির্ভরযোগ্যতার সঠিক মান নির্ধারণ করা যায় না।

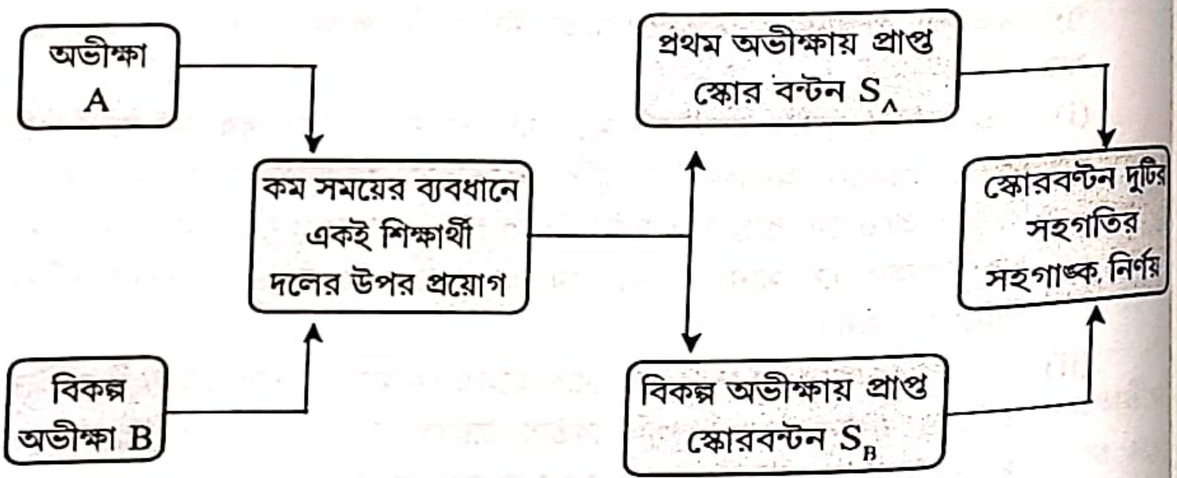
- **কার্যকরী করার উপায় ঃ**

অভীক্ষার প্রথম ও দ্বিতীয় প্রয়োগের মধ্যে যথার্থ সময়ের ব্যবধান রাখা উচিত (3–4 সপ্তাহ) যাতে স্মৃতিশক্তি, অনুশীলন, অভিজ্ঞতা সঞ্চালন, বিস্মৃতি নির্ভরযোগ্যতা সহগাঙ্ককে প্রভাবিত করতে না পারে।

Scanned with CamScanner

২.৩.১.২.২ বিকল্প পদ্ধতি বা সমান্তরাল পদ্ধতি বা সমতুল্য অভীক্ষা পদ্ধতি (Alternate Form or Parallel Form or Equivalent Test Method) :

- **পদ্ধতি :** এই পদ্ধতিতে কোনো অভীক্ষার নির্ভরযোগ্যতা নির্ণয়ের জন্য একই সঙ্গে দুটি ভিন্ন (different) কিন্তু সদৃশ (equivalent) অভীক্ষা গঠন করা হয়। অভীক্ষাদুটির বিষয়বস্তু, উদ্দেশ্য, কাঠিন্যমাত্রা, নির্দেশাবলির প্রকৃতি, অভীক্ষাপদের পার্থক্য নির্ণায়ক মান, সময়সীমা ইত্যাদির মধ্যে পরিপূর্ণ সাদৃশ্য থাকে। তাই অভীক্ষাদুটিকে "সমান্তরাল অভীক্ষা" বলা হয়। অভীক্ষাদুটিকে খুব কম সময়ের ব্যবধানে একদল নির্বাচিত শিক্ষার্থীর উপর পর পর প্রয়োগ করা হয় এবং প্রাপ্ত স্কোরবণ্টন দুটির সহগতির সহগাঙ্ক (Correlation Co-efficient) নির্ণয় করা হয়। এই সহগাঙ্ক অভীক্ষা দুটির তুল্যতার পরিমাপ (measure of equivalence) যা শিক্ষার্থীর কোনো বৈশিষ্ট্য পরিমাপে অভীক্ষা দুটির সাদৃশ্যতার মাত্রা নির্দেশ করে।



- **উদাহরণ :** 1937 সালের বুদ্ধির অভীক্ষার স্ট্যানফোর্ড বিনে সংস্করণে (1937, Standford-Binet Scale) 'L' ফর্ম এবং 'M' ফর্ম নামক দুটি সমান্তরাল অভীক্ষা গঠন করে নির্ভরযোগ্যতা নির্ণয় করা হয়েছিল। অভীক্ষা দুটিতে একই পদ্ধতিতে আদর্শায়িত করা হয়েছিল। 7 বছর বয়সী শিক্ষার্থীদের একটি নমুনা দলের উপর অভীক্ষা দুটি প্রয়োগ করে দেখা গেছে ফর্ম দুটির সহগাঙ্ক 0.91।

- **সমান্তরাল অভীক্ষার বৈশিষ্ট্য :** অভীক্ষা দুটির মধ্যে নিম্নলিখিত বৈশিষ্ট্যগুলি থাকলে তাদের সমান্তরাল অভীক্ষা বলা হবে। যথা—
  (i) উভয়ক্ষেত্রে অভীক্ষাপদের সংখ্যা একই হবে।
  (ii) অভীক্ষাপদের প্রকৃতি, কাঠিন্যমাত্রা ও বিন্যাসে সাদৃশ্য থাকবে।

Scanned with CamScanner

(iii) উভয় অভীক্ষার সময়সীমা একই হতে হবে।
(iv) বিষয়বস্তুগত ও উদ্দেশ্যগত সাদৃশ্য থাকবে।
(v) উভয় অভীক্ষার উত্তরদানের জন্য একইরকম নির্দেশাবলি থাকবে।
(vi) অভীক্ষা পরিচালনা ও নম্বরদানের ক্ষেত্রে একই নিয়ম ও নির্দেশ অনুসরণ করা হবে।

- **ত্রুটি ঃ** এই পদ্ধতিতে নিম্নলিখিত ত্রুটিগুলি লক্ষ করা যায়—

(i) অভীক্ষাদুটি সদৃশ হওয়ায় প্রথম অভীক্ষাগ্রহণের অভিজ্ঞতা বিকল্প অভীক্ষার ফলাফলকে প্রভাবিত করে।
(ii) সমান্তরাল অভীক্ষা গঠনে ব্যক্তিগত প্রভাব পড়ে। কারণ, দুটি অভীক্ষাপদ ভিন্ন ব্যক্তির কাছে সমকাঠিন্যমান সম্পন্ন মনে নাও হতে পারে।
(iii) দুটি অভীক্ষাগ্রহণের মধ্যে সময়ের ব্যবধান কম হলে নির্ভরযোগ্যতার মান বেশি হবে। আবার সময়ের ব্যবধান বেশি হলে নির্ভরযোগ্যতার মান কম হয়।
(iv) দুটি যথাযথ সমান্তরাল অভীক্ষা গঠন সময়সাপেক্ষ ও পরিশ্রমসাধ্য।

উপরোক্ত ত্রুটিগুলি থাকা সত্ত্বেও সমান্তরাল অভীক্ষা পদ্ধতি হল নির্ভরযোগ্যতা নির্ধারণের অধিকতর সন্তোষজনক কৌশল। অভীক্ষাদুটি প্রয়োগের মধ্যে কমপক্ষে 2–4 সপ্তাহ সময়ের ব্যবধান রাখা হলে অভীক্ষার্থীর পূর্ব অভিজ্ঞতার পুনরাবৃত্তি হওয়ার সুযোগ হ্রাস পায়।

▧ ২.৩.১.২.৩ অর্ধ-দ্বিখণ্ডিত পদ্ধতি (Split-half Method or Intra Split-half Method) ঃ

পুনঃ অভীক্ষা পদ্ধতি ও সমান্তরাল অভীক্ষা পদ্ধতিতে দুবার অভীক্ষা প্রয়োগের মধ্যে সময়ের ব্যবধান থাকায় অভীক্ষার্থীর স্মৃতি ও অভিজ্ঞতা সঞ্চালন নির্ভরযোগ্যতা নির্ণয়ে প্রভাব ফেলে। ফলে নির্ভরযোগ্যতার স্থায়ী মান নির্ণয় করা যায় না। অর্ধ-দ্বিখণ্ডিত পদ্ধতিতে এই অসুবিধা দূর করার চেষ্টা করা হয়েছে।

- **পদ্ধতি ঃ**

এই পদ্ধতিতে কোনো অভীক্ষার নির্ভরযোগ্যতার সহগাঙ্ক নির্ণয় করতে হলে নিম্নলিখিত পর্যায়গুলি অনুসরণ করতে হয়।

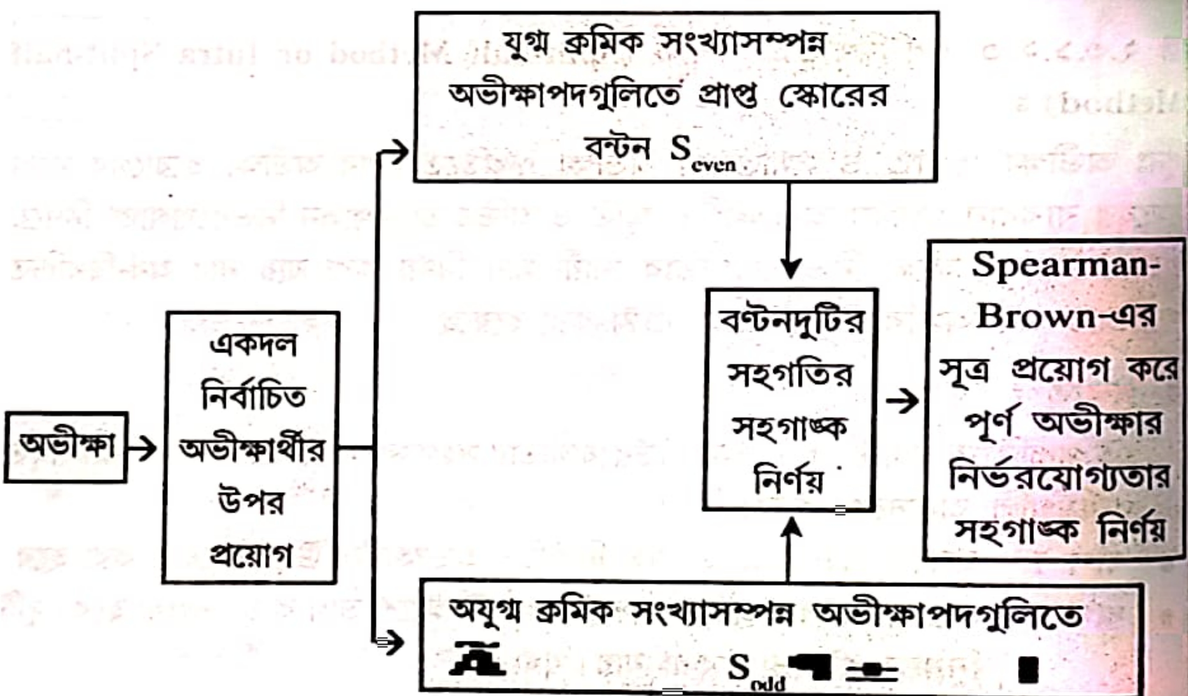
- **ধাপ 1** প্রথমে অভীক্ষাটি একদল নির্বাচিত অভীক্ষার্থীর উপর প্রয়োগ করা হবে।
- **ধাপ 2** প্রয়োগের পর অভীক্ষাপদগুলিকে দুটি অংশে ভাগ করে নেওয়া হবে। দুটি নিয়মে এই ভাগ করা যায়। যথা—

Scanned with CamScanner

(i) **যুগ্ম-অযুগ্ম (even-odd) :** কাঠিন্যতার ক্রমানুসারে সাজিয়ে যুগ্ম ক্রমিক সংখ্যাসম্পন্ন অভীক্ষাপদগুলিকে একখণ্ডে এবং অযুগ্ম ক্রমিক সংখ্যাসম্পন্ন পদগুলিকে অপর খণ্ডে নেওয়া হয়। অর্থাৎ প্রথম খণ্ডে 2 নং, 4 নং, 6 নং, ... পদগুলি ও দ্বিতীয় খণ্ডে 1 নং, 3 নং, 5 নং.......... পদগুলি থাকবে।

(ii) **প্রথম অর্ধ - দ্বিতীয় অর্ধ (1st Half-2nd Half) :** সমস্ত অভীক্ষাপদগুলির প্রথম অর্ধ এক অংশে ও দ্বিতীয় অর্ধ অপর অংশে নেওয়া হয়। তবে এমনভাবে নিতে হবে যেন অর্ধদুটি পরস্পরের সমান্তরাল বা বিকল্প হয়। ধরা যাক, 30 টি অভীক্ষাপদের 1-15 নং পদ প্রথম অংশে ও 16-30 নং পদ দ্বিতীয় অংশে নেওয়া হল।

- **ধাপ 3** অভীক্ষার অন্তর্গত যুগ্ম ক্রমিক সংখ্যাসম্পন্ন অভীক্ষাপদগুলিতে প্রাপ্ত স্কোর এবং অযুগ্ম ক্রমিক সংখ্যা সম্পন্ন অভীক্ষাপদগুলিতে প্রাপ্ত স্কোর-এর পৃথক দুটি বণ্টন নির্ণয় করা হবে।
- **ধাপ 4** দুটি পৃথক স্কোর বণ্টনের সহগতির সহগাঙ্ক নির্ণয় করা হবে। যা অভীক্ষার দুটি অর্ধের নির্ভরযোগ্যতার সহগাঙ্ক (Reliability Co-efficient of Half-tests)।
- **ধাপ 5** খণ্ডিত অংশের নির্ভরযোগ্যতার মান থেকে পূর্ণ অভীক্ষার নির্ভরযোগ্যতা নির্ণয়ের জন্য স্পিয়ারম্যান ও ব্রাউনের (Spearman–Brown Formula) সূত্র প্রয়োগ করা হবে।



Scanned with CamScanner

• স্পিয়ারম্যান-ব্রাউনের সূত্র (Spearman-Brown prophecy formula) :

(i) অভীক্ষাটি দুটি খণ্ডে বিভক্ত হলে : $$r_t = \frac{2r_{1/2}}{1+r_{1/2}}$$

এখানে,

r_t = পূর্ণ অভীক্ষার নির্ভরযোগ্যতার সহগাঙ্ক (full test reliability)।

$r_{1/2}$ = অর্ধ-অভীক্ষাদুটির সহগতির সহগাঙ্ক (half-test reliability)।

ধরা যাক, অর্ধ-অভীক্ষার সহগতির সহগাঙ্ক 0.50 তাহলে পূর্ণ-অভীক্ষার নির্ভরযোগ্যতার সহগাঙ্ক হবে :

$$r_t = \frac{2\times0.50}{1+0.50} = \frac{1}{1.50} \approx 0.67$$

(ii) অভীক্ষার দৈর্ঘ্য n গুণ বৃদ্ধি করলে বা nটি সমান্তরাল অভীক্ষা প্রয়োগ করলে :

$$r_{nn} = \frac{nr_t}{1+(n-1)r_t}$$

r_{nn} = অভীক্ষার দৈর্ঘ্য বৃদ্ধির ফলে নির্ভরযোগ্যতার পরিবর্তিত মান।

n = অভীক্ষা খণ্ডের সংখ্যা বা সমান্তরাল অভীক্ষার সংখ্যা বা অভীক্ষাপদের সংখ্যা যে গুণিতক হারে বৃদ্ধি করা হয়েছে।

r_t = মূল অভীক্ষার নির্ভরযোগ্যতা।

উদাহরণ :

(i) ধরা যাক 50 টি অভীক্ষাপদ সম্পন্ন কোনো অভীক্ষাকে 25 টি যুগ্ম-ক্রমিক সংখ্যা সম্পন্ন ও 25 টি অযুগ্ম-ক্রমিক সংখ্যাসম্পন্ন দুটি অর্ধে ভাগ করা হল। খণ্ডদুটির সহগতির সহগাঙ্ক 0.50 হলে পূর্ণ অভীক্ষার নির্ভরযোগ্যতা

$$r = \frac{nr_t}{1+(n-1)r_t} = \frac{2\times0.50}{1+(2-1)\times0.50} = \frac{1.00}{1.50} \approx 0.67$$

(ii) যদি অভীক্ষাপদের সংখ্যা দ্বিগুণ করা হয় অর্থাৎ, 100 টি প্রশ্ন থাকে তবে খণ্ডের সংখ্যা $n = \frac{100}{25} = 4$

Scanned with CamScanner

$$\therefore r = \frac{4 \times 0.50}{1 + (4-1) \times 0.50} = \frac{2.00}{2.50} = 0.80$$

সুতরাং, অভীক্ষার দৈর্ঘ্য বৃদ্ধি করলে নির্ভরযোগ্যতার মান বৃদ্ধি পায়।

- **পদ্ধতির সুবিধা :**

পুনঃঅভীক্ষা পদ্ধতি ও সমান্তরাল অভীক্ষা পদ্ধতির চেয়ে এই পদ্ধতি অধিক সুবিধাজনক। কারণ—

(i) বিবেচ্য অভীক্ষাটিকে একবারই প্রয়োগ করতে হয়। তাই নির্ভরযোগ্যতার মান নির্ণয়ে অভীক্ষার্থীর স্মৃতি ও অভিজ্ঞতা সঞ্চালনের কোনো প্রভাব পড়ে না।

(ii) দুটি পৃথক (বিকল্প) অভীক্ষা গঠন করার প্রয়োজন না থাকায় সময়ও কম লাগে আবার পরিশ্রমও লাঘব হয়।

- **পদ্ধতির ত্রুটি :**

এই পদ্ধতিতে কোনো অভীক্ষার নির্ভরযোগ্যতার যে মান নির্ণয় করা হয় তা সম্পূর্ণভাবে ত্রুটিমুক্ত বলা যায় না। কারণ—

(i) অভীক্ষাটিকে যে দুটি খণ্ডে বিভক্ত করা হয়েছে তাদের মধ্যেকার সাদৃশ্য সম্পূর্ণভাবেই অনুমানভিত্তিক। তাই নির্ভরযোগ্যতার যে মান পাওয়া যায় তা অনুমানের উপর নির্ভরশীল।

(ii) অভীক্ষার অন্তর্গত অভীক্ষাপদগুলির ভিন্ন ভিন্ন বিন্যাসের দরুন অভীক্ষার খণ্ডিতাংশ দুটির মধ্যে পৃথক পৃথক সহগতির সহগাঙ্কের মান পাওয়া যাবে। ফলে নির্ভরযোগ্যতার স্থায়ী মান পাওয়া যাবে না।

⊠ ২.৩.১.২.৪ অন্তপদীয় সামঞ্জস্য পদ্ধতি (Inter-item Consistency Method):

এই পদ্ধতিতে কোনো অভীক্ষার নির্ভরযোগ্যতার মান নির্ধারণ করতে হলে অভীক্ষাটিকে দ্বিখণ্ডিত করার প্রয়োজন হয় না। অর্ধ-দ্বিখণ্ডিত পদ্ধতির মতো এই পদ্ধতিতেও অভীক্ষার অন্তর্গত অভীক্ষাপদগুলির অভ্যন্তরীণ সামঞ্জস্যতা নির্ণয় (Measure of internal consistency) করা হয়।

- **পদ্ধতি :**

মনোবিদ্ কুডার ও রিচার্ডসন (Kuder-Richardson) এই পদ্ধতিতে কোনো অভীক্ষার নির্ভরযোগ্যতা নির্ণয়ের দুটি সূত্র গঠন করেছেন। যথা—

(i) কুডার-রিচার্ডসন সূত্র-20 (Kuder-Richardson Formula-20 or KR-20)

Scanned with CamScanner

(ii) কুডার-রিচার্ডসন সূত্র-21 (Kudar-Richardson Formula-21 or KR-21)

(i) KR-20 ঃ সূত্রটি হল—

$$r_t = \frac{n}{(n-1)} \times \frac{\sigma_t^2 - \sum pq}{\sigma_t^2}$$

এখানে,

r_t = বিবেচ্য অভীক্ষাটির নির্ভরযোগ্যতার সহগাঙ্ক (Reliability Co-efficient of the whole test) ।

n = অভীক্ষার অন্তর্গত অভীক্ষাপদের সংখ্যা (Number of test items in the test)।

σ_t = প্রাপ্ত স্কোরবণ্টনের সম্যক চ্যুতি (Standard deviation of the test scores)।

p = কোনো অভীক্ষাপদের সঠিক উত্তরদাতার অনুপাত (Proportion of the group answering a test item correctly)।

q = কোনো অভীক্ষাপদের ভুল উত্তরদাতার অনুপাত (Proportion of the group answering a test item incorrrectly)।

এই সূত্র প্রয়োগ করে কোনো অভীক্ষার নির্ভরযোগ্যতার সহগাঙ্ক নির্ণয়ের জন্য নিম্নলিখিত পর্যায়গুলি অনুসরণ করতে হয়—

- **ধাপ 1** নির্বাচিত একদল অভীক্ষার্থীর উপর বিবেচ্য অভীক্ষাটি একবার প্রয়োগ করা হয়। অভীক্ষায় প্রাপ্ত স্কোরবণ্টনটির সম্যক চ্যুতি σ_t নির্ণয় করা হবে।

- **ধাপ 2** প্রতিটি অভীক্ষাপদের জন্য, মোট অভীক্ষার্থীদের যত অংশ (p) সঠিক উত্তর দিয়েছে এবং যত অংশ (q) ভুল উত্তর দিয়েছে তা নির্ণয় করা হবে। যদি N সংখ্যক অভীক্ষার্থীর মধ্যে P সংখ্যক অভীক্ষার্থী সঠিক উত্তর দেয় এবং Q সংখ্যক অভীক্ষার্থী ভুল উত্তর দেয় তবে

$p = \left(\frac{P}{N}\right)$ এবং $q = \left(\frac{Q}{N}\right)$.

Scanned with CamScanner

- **ধাপ 3** প্রতিটি অভীক্ষাপদের জন্য p এবং q এর গুণফল নির্ণয় করে সমস্ত অভীক্ষাপদগুলির pq-এর সমষ্টি অর্থাৎ, $\sum pq$ নির্ণয় করা হবে।
- **ধাপ 4** KR–20 সূত্রে মোট অভীক্ষাপদের সংখ্যা (n), সম্যকচ্যুতি (σ_t) ও $\sum pq$ -এর মান বসিয়ে অভীক্ষার নির্ভরযোগ্যতার সহগাঙ্ক নির্ণয় করা হবে।

এই সূত্র প্রয়োগের ক্ষেত্রে অপেক্ষাকৃত জটিল গণনার প্রয়োজন হয়। এই সূত্রের বিকল্প হিসাবে এবং অপেক্ষাকৃত সহজে অভীক্ষার নির্ভরযোগ্যতা-সহগাঙ্ক নির্ণয়ের জন্য কুডার-রিচার্ডসন আর একটি সূত্র গঠন করেছেন যা কুডার-রিচার্ডসন সূত্র 21 (KR–21) নামে পরিচিত।

(ii) KR-21 ঃ সূত্রটি হল—

এখানে,

r_t = বিবেচ্য অভীক্ষাটির নির্ভরযোগ্যতার সহগাঙ্ক (Reliability Co-efficient of the whole test)।

n = অভীক্ষার অন্তর্গত অভীক্ষাপদের সংখ্যা (Number of test items in the test)।

M = প্রাপ্ত স্কোরবণ্টনের গড় (Arithmetic mean or mean of the test scores)।

σ_t = প্রাপ্ত স্কোরবণ্টনের সম্যক চ্যুতি (Standard deviation of the test scores)।

এই সূত্র প্রয়োগ করে কোনো অভীক্ষার নির্ভরযোগ্যতার সহগাঙ্ক নির্ণয়ের জন্য নিম্নলিখিত পর্যায়গুলি অনুসরণ করতে হয়—

- **ধাপ 1** নির্বাচিত একদল অভীক্ষার্থীর উপর বিবেচ্য অভীক্ষাটি একবার প্রয়োগ করা হবে। অভীক্ষায় প্রাপ্ত স্কোরবণ্টনটি নির্ণয় করতে হবে।
- **ধাপ 2** প্রাপ্ত স্কোরবণ্টনটির গড় (M) ও সম্যকচ্যুতি (σ_t) নির্ণয় করা হবে।
- **ধাপ 3** KR-21 সূত্রে মোট অভীক্ষাপদের সংখ্যা (n), স্কোরবণ্টনের গড় (M) ও সম্যকচ্যুতির (σ_t) মান বসিয়ে অভীক্ষাটির নির্ভরযোগ্যতার সহগাঙ্ক নির্ণয় করা হবে।

Scanned with CamScanner

- **পদ্ধতির সুবিধা :**

এই পদ্ধতির সুবিধাগুলি হল—

(i) এই পদ্ধতিতে অভীক্ষাটি পুনঃপ্রয়োগের প্রয়োজন হয় না, মাত্র একবারই প্রয়োগ করা হয়।

(ii) অর্ধ-দ্বিখণ্ডিত পদ্ধতির (Split-half Method) মতো এক্ষেত্রে বাস্তবে পূর্ণ অভীক্ষাটিকে দ্বিখণ্ডিত করার প্রয়োজন হয় না।

- **পদ্ধতির ত্রুটি :**

এই পদ্ধতিতে অভীক্ষার নির্ভরযোগ্যতার যে মান পাওয়া যায় তা অভীক্ষাপদগুলির প্রকৃতির উপর নির্ভরশীল। অভীক্ষাপদগুলি যত সমজাতীয় (Homogeneous) হবে নির্ভরযোগ্যতার মান তত সন্তোষজনক হবে। ধরা যাক, একদল নির্বাচিত শিক্ষার্থীর উপর দুটি অভীক্ষা প্রয়োগ করা হল। যার একটিতে শুধু গাণিতিক 'যোগ' (addition) সংক্রান্ত অভীক্ষাপদ এবং অপরটিতে বিয়োগ (subtraction), গুণ (multiplicatioin) ও ভাগ (division) সংক্রান্ত অভীক্ষাপদ রয়েছে। স্বভাবতই প্রথম অভীক্ষাটির অন্তর্গত অভীক্ষাপদগুলি সমজাতীয় হওয়ায় অভীক্ষাটির অন্তপদীয় সামঞ্জস্যতা (Inter-item consistency) অপেক্ষাকৃত বেশি। দ্বিতীয় অভীক্ষার অন্তর্গত অভীক্ষাপদগুলি বিষমপ্রকৃতির (Heterogeneous) হওয়ায় বিভিন্ন অভীক্ষাপদে অভীক্ষার্থীদের উত্তরদানে কোনো সহগতি থাকে না বললেই চলে। ফলে নির্ভরযোগ্যতার পরিমাপ ত্রুটিপূর্ণ হয়।

- **উদাহরণ :** একটি অভীক্ষার পদসংখ্যা 15। প্রত্যেক সঠিক উত্তরদানে 1 ও ভুল উত্তরদানে 0 স্কোর ধার্য করা হয়েছে। অভীক্ষাটি 50 জন অভীক্ষার্থীর একটি নির্বাচিত দলের উপর প্রয়োগ করা হলো। প্রতিটি অভীক্ষাপদের সঠিক উত্তরদাতার সংখ্যার একটি বণ্টন নীচে দেওয়া হল। অভীক্ষার্থীদের প্রাপ্ত মোট স্কোরের গড় 9 ও সমকচ্যুতি 4 হলে KR_{20} ও KR_{21} সূত্র পৃথকভাবে প্রয়োগ করে অভীক্ষাটির নির্ভরযোগ্যতার সহগাঙ্ক নির্ণয় করা হল।

অভীক্ষাপদের ক্রমিক নং	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
অভীক্ষাপদটির সঠিক উত্তর-দাতার সংখ্যা	21	25	35	44	27	31	32	36	42	25	35	30	20	48	40

Scanned with CamScanner

সমাধান ঃ

এখানে, অভীক্ষাপদের সংখ্য $n = 15$
মোট অভীক্ষার্থীর সংখ্যা $N = 50$
প্রাপ্ত স্কোরের গড় $M = 9$
সম্যকচ্যুতি $\sigma_t = 4$

KR–20 :

অভীক্ষাপদের ক্রমিক নং	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$P = Np$	21	25	35	44	27	31	32	36	42	25	35	30	20	48	40
$Q = Nq$	29	25	15	6	23	19	18	14	8	25	15	20	30	20	10
$PQ = N^2pq$	609	625	525	264	621	589	576	504	336	625	525	600	600	96	400
													$\Sigma PQ = 7495$		

কুডার-রিচার্ডসন সূত্র-20 অনুসারে অভীক্ষাটির নির্ভরযোগ্যতার সহগাঙ্ক হল-

$$r_t = \frac{n}{(n-1)} \times \frac{\sigma_t^2 - \sum pq}{\sigma_t^2}$$

$$= \frac{15}{(15-1)} \times \frac{(4)^2 - 2.998}{(4)^2}$$

$$= \frac{15}{14} \times \frac{16 - 2.998}{16}$$

$$= \frac{15}{14} \times \frac{13.002}{16} \approx 0.87$$

$\therefore\ r_t \approx 0{\cdot}87$

$$\Sigma PQ = \Sigma N^2pq = \Sigma N^2\ pq$$

$$\therefore\ N^2 \Sigma pq = 7495$$

$$\therefore\ \Sigma pq = \frac{7495}{N^2}$$

$$= \frac{7495}{(50)^2}$$

$$= 2.998$$

KR– 21 :

"কুডার-রিচার্ডসন সূত্র-21" অনুসারে অভীক্ষাটির নির্ভরযোগ্যতার সহগাঙ্ক হল—

$$r_t = \frac{n}{(n-1)} \times \left\{1 - \frac{M(n-M)}{n\sigma_t^2}\right\}$$

$$= \frac{15}{(15-1)} \times \left\{1 - \frac{9(15-9)}{15\times(4)^2}\right\}$$

$$= \frac{15}{14} \times \left\{1 - \frac{9\times6}{15\times16}\right\}$$

$$= \frac{15}{14} \times \left\{1 - \frac{9}{40}\right\} = \frac{15}{14} \times \frac{31}{40} \approx 0.83 \qquad \therefore\ r_t \approx 0{\cdot}83$$

❖ ২.৩.১.৩ পরিমাপের ত্রুটি ও নির্ভরযোগ্যতা (Errors of Measurement and Reliability) ঃ

কোনো শিক্ষাগত বা মানসিক বৈশিষ্ট্য পরিমাপের অভীক্ষায় শিক্ষার্থীর প্রাপ্ত স্কোর (obtained score) তার প্রকৃত পারদর্শিতা বা সামর্থ্যের সূচক নয়। কারণ, স্কোরের মধ্যে পরিমাপজনিত কিছু ত্রুটি থেকে যায়।

মনোবিদ **H.E. Garrett**-এর মতে "A score on a mental test may be thought of as an index of the examinee's 'true ability' plus errors of measurement."

যদি স্কোরের মধ্যে ত্রুটির পরিমাণ কম থাকে তাহলে অভীক্ষার নির্ভরযোগ্যতার মান উচ্চ হবে, বিপরীতক্রমে, যদি স্কোরের মধ্যে ত্রুটির পরিমাণ বেশি হয় তাহলে অভীক্ষার নির্ভরযোগ্যতার মান খুব কম হবে।

কোনো অভীক্ষায় একজন শিক্ষার্থীর প্রাপ্ত স্কোর (obtained score), প্রকৃত স্কোর (true score) এবং ত্রুটি (error)-এর গাণিতিক সম্পর্কটি হল ঃ

প্রাপ্ত স্কোর = প্রকৃত স্কোর + ত্রুটি

ধরা যাক, X_0 = কোনো অভীক্ষায় একজন শিক্ষার্থীর প্রাপ্ত স্কোর (obtained score)

X_T = ঐ শিক্ষার্থীর প্রকৃত স্কোর (true score)

X_e = সম্ভাব্য ত্রুটি বা পরিবর্তনশীল ত্রুটি (chance error or variable error)

$\therefore\ X_0 = X_T + X_e$ .... (i)

Scanned with CamScanner

∴ ভ্যারিয়েন্সের ধর্ম ব্যবহার করে পাই

$Var(X_0) = Var(X_T + X_e) = Var(X_T) + Var(X_e) + 2\,cov(X_T, X_e)$ ......

$\therefore \sigma_0^2 = \sigma_T^2 + \sigma_e^2 + 2r_{Te}\sigma_T\sigma_e$ ....(iii)

সম্ভাব্য ত্রুটি ও প্রকৃত স্কোর এর মধ্যে কোনো সহগতি না থাকলে

অর্থাৎ, $r_{Te} = 0$ হলে

$\sigma_0^2 = \sigma_T^2 + \sigma_e^2$ .... (iv)

(iv) নং সমীকরণের উভয় দিককে σ_0^2 দিয়ে ভাগ করে পাওয়া যায়—

$$1 = \frac{\sigma_T^2}{\sigma_0^2} + \frac{\sigma_e^2}{\sigma_0^2} \text{ ....(v)}$$

$\left(\frac{\sigma_T^2}{\sigma_0^2}\right)$ এই অনুপাতটির মান বড়ো হলে নির্ভরযোগ্যতার মান উচ্চ হ

এবং $\left(\frac{\sigma_e^2}{\sigma_0^2}\right)$ এর মান বড়ো হলে নির্ভরযোগ্যতার মান নিম্ন হবে।

প্রকৃত স্কোর ও সম্ভাব্য ত্রুটি পরস্পরের উপর নির্ভরশীল না হলে বা তাদে মধ্যে কোনো সহগতি নেই ধরা হলে, আমরা লিখতে পারি—

অভীক্ষার নির্ভরযোগ্যতার সহগাঙ্ক ঃ $r_t = \frac{\sigma_T^2}{\sigma_0^2} = \frac{\text{true variance}}{\text{test variance}}$ .... (vi)

সমীকরণ (v) ও (vi) থেকে পাই

$$r_t = \left(1 - \frac{\sigma_e^2}{\sigma_0^2}\right) \text{ ....(vii)}$$

সুতরাং, নির্ভরযোগ্যতার সহগাঙ্ক হল কোনো পরিমাপের নির্ভুলতার সূচ (Reliability Coefficient is an index of precision of measurement)। য কোনো অভীক্ষার নির্ভরযোগ্যতা-সহগাঙ্কের মান 0.90 হয় তবে প্রাপ্ত ভ্যারিয়েন্স (variance of test score) 90% হল প্রকৃত-স্কোর ভ্যারিয়েন্স (true-score variance) এবং 10% হল সম্ভাব্য ত্রুটির ভ্যারিয়েন্স (Chance error variance)।

- **প্রাপ্ত স্কোরের আদর্শ ত্রুটি বা পরিমাপের আদর্শ ত্রুটি (The Standard error of an obtained score or the standard error of measurement) ঃ**

সমীকরণ (vii) থেকে আমরা লিখতে পারি—

Scanned with CamScanner

$$\frac{\sigma_e^{\,2}}{\sigma_0^{\,2}} = (1 - r_t)$$

বা, $\sigma_e = \sigma_0 \sqrt{1 - r_t}$

এখানে,

σ_e = প্রাপ্ত স্কোরের আদর্শ ত্রুটি বা পরিমাপের আদর্শ ত্রুটি (the SE of an obtained score)।

σ_0 = অভীক্ষায় প্রাপ্ত স্কোর বণ্টনের সম্যকচ্যুতি (the standard deviation of test scores)।

r_1 = অভীক্ষার নির্ভরযোগ্যতা-সহগাঙ্ক (the reliability coefficient of the test)।

❖ ২.৩.১.৪ অভীক্ষার নির্ভরযোগ্যতা হ্রাসের কারণসমূহ (Causes of low reliability of a test) :

শিক্ষামনোবিদ্‌গণ কোনো অভীক্ষার নির্ভরযোগ্যতা হ্রাসের কারণসমূহকে তিনটি শ্রেণিতে বিভক্ত করেছেন। যথা—

(1) অভীক্ষা নির্ভর কারণ।
(2) অভীক্ষার্থী নির্ভর কারণ।
(3) অভীক্ষক নির্ভর কারণ।

উপরোক্ত কারণগুলি নিম্নে সংক্ষেপে আলোচনা করা হল।

(1) অভীক্ষা নির্ভর কারণ :

কোনো অভীক্ষার নিম্নলিখিত সাংগঠনিক ত্রুটিগুলি থাকলে অভীক্ষার নির্ভরযোগ্যতা হ্রাস পায়। যথা—

(i) **অভীক্ষার দৈর্ঘ্য (Length of a test) :** কোনো অভীক্ষার দৈর্ঘ্য বৃদ্ধি করলে তার নির্ভরযোগ্যতা বৃদ্ধি পায় এবং অভীক্ষাটি সংক্ষিপ্ত হলে অর্থাৎ অভীক্ষায় পর্যাপ্ত সংখ্যক অভীক্ষাপদ না থাকলে অভীক্ষার নির্ভরযোগ্যতা হ্রাস পায়।

(ii) **অভীক্ষাপদের কাঠিন্যমান (Difficulty value of test items) :** কোনো অভীক্ষার অন্তর্গত সব অভীক্ষাপদগুলি খুবই সহজ বা খুবই কঠিন হলে অর্থাৎ, বিভিন্ন কাঠিন্যমানের অভীক্ষাপদ নির্বাচন করতে না পারলে অভীক্ষাটির নির্ভরযোগ্যতা হ্রাস পাবে।

Scanned with CamScanner

(iii) **অভীক্ষাপদগুলির উদ্দেশ্যহীনতা (Lack of objectives of the test items)** : শিক্ষার্থীর যে শিক্ষাগত যোগ্যতা বা মানসিক বৈশিষ্ট্য পরিমাপের উদ্দেশ্যে অভীক্ষাটি গঠন করা হয়েছে তার সঙ্গে সঙ্গতি না রেখে অভীক্ষাপদ নির্বাচন করলে অভীক্ষাটির নির্ভরযোগ্যতা হ্রাস পায়।

(iv) **অভীক্ষাপদের বিক্ষিপ্ত বিন্যাস (Haphazard arrangement of test items)** : অভীক্ষাপদগুলি কাঠিন্যতার মান অনুসারে সহজ থকে কঠিন এই ক্রমে যথাযথভাবে বিন্যস্ত না করা থাকলে অভীক্ষার্থীরা উত্তরদানে সমস্যার সম্মুখীন হয়। ফলে অভীক্ষার নির্ভরযোগ্যতা হ্রাস পায়।

(v) **অভীক্ষাপদের পুনরাবৃত্তি ও পরস্পর নির্ভরশীলতা (Repetition and inter-dependence of test items)** : পরস্পর নির্ভরশীল অভীক্ষাপদ কিংবা একই জাতীয় অভীক্ষাপদের পুনরাবৃত্তি উভয়ই অভীক্ষার নির্ভরযোগ্যতা হ্রাসের কারণ। কারণ, অভীক্ষার্থীরা কোনো একটি অভীক্ষাপদের সঠিক উত্তর দিতে পারলে সমজাতীয় অন্য অভীক্ষাপদের উত্তর সেখান থেকেই পেয়ে যায়।

(vi) **অনুমান-সহায়ক অভীক্ষাপদ (Test items that can be answered by guessing)** : কোনো অভীক্ষায় সত্য-মিথ্যা জাতীয় কিংবা বহুমুখী নির্বাচনধর্মী অভীক্ষাপদের উত্তরদানে শিক্ষার্থীদের মধ্যে অনুমানপ্রবণতা লক্ষ করা যায় যা অভীক্ষার নির্ভরযোগ্যতা হ্রাসের কারণ। তাই অভীক্ষাপদগুলি এমনভাবে নির্বাচন করতে হবে যেন তারা অনুমানভিত্তিক উত্তর দিতে না পারে।

- [বিঃ দ্রঃ—নৈর্ব্যক্তিক অভীক্ষার নির্ভরযোগ্যতা বজায় রাখার জন্য একটা সংশোধনী ব্যবহার করা হয়। সূত্রটি হল :

$$S = \left[R - \frac{W}{(n-1)}\right]$$

যেখানে,

S = প্রকৃত স্কোর (True Score)
R = সঠিক উত্তরদানের সংখ্যা (Number of right responses)
W = ভুল উত্তরদানের সংখ্যা (Number of wrong responses)।
n = অভীক্ষাপদে প্রদত্ত বিকল্প উত্তরের সংখ্যা (Number of alternatives provided with test item)।

ধরা যাক, কোনো অভীক্ষার্থী 3টি বিকল্প সম্ভাব্য উত্তর সম্বলিত 10টি বহুমুখী-নির্বাচনধর্মী অভীক্ষাপদের (Multiple-choice type questions) মধ্যে 6টি অভীক্ষাপদের সঠিক

Scanned with CamScanner

উত্তর দিয়েছে। প্রতিটি সঠিক উত্তরের জন্য 1 স্কোর ধার্য করা হলে, তার প্রকৃত স্কোর—

$$S = R - \frac{W}{(n-1)} = 6 - \frac{4}{(3-1)} = 6 - \frac{4}{2} = 6 - 2 = 4$$

(R = 6, W = (10 − 6) = 4, n = 3)]

(2) **অভীক্ষার্থী নির্ভর কারণ :**

অভীক্ষার্থীর দৈহিক, মানসিক ও প্রাক্ষোভিক অবস্থা অভীক্ষার নির্ভরযোগ্যতা হ্রাসের অন্যতম কারণ। যেমন—

(i) **শারীরিক অবস্থা (Physical Condition) :** অভীক্ষা প্রয়োগের সময় অভীক্ষার্থী কোনো রকম শারীরিক অসুস্থতা বা অস্বাচ্ছন্দ্য বোধ করলে সঠিকভাবে প্রতিক্রিয়া করতে পারে না। ফলে অভীক্ষার নির্ভরযোগ্যতার সঠিক পরিমাপ পাওয়া যায় না।

(ii) **মানসিক প্রস্তুতি (Mental Readiness) :** অভীক্ষা গ্রহণের মতো মানসিক প্রস্তুতি যদি অভীক্ষার্থীর না থাকে তাহলে অভীক্ষার নির্ভরযোগ্যতা হ্রাস পায়।

(iii) **প্রাক্ষোভিক অবস্থা (Emotional Condition) :** অভীক্ষা গ্রহণের সময় অভীক্ষার্থীরা তাদের প্রাক্ষোভিক অবস্থাকে নিয়ন্ত্রণ করতে না পারলে সঠিক উত্তরদানে ব্যর্থ হয়। ফলে অভাক্ষার নির্ভরযোগ্যতা হ্রাস পায়।

(iv) **পারিপার্শ্বিক অবস্থা (External Situation) :** অভীক্ষাগ্রহণের স্থান ও পারিপার্শ্বিক অবস্থা অনুকূল না হলে, অভীক্ষা গ্রহণের সময় অভীক্ষার্থীকে সাহায্য করলে কিংবা অভীক্ষার্থীরা উত্তরদানে অসৎ উপায় অবলম্বন করলে অভীক্ষার নির্ভরযোগ্যতা হ্রাস পায়।

(v) **অভিজ্ঞতা সঞ্চালন (Transfer of Experience) :** অভীক্ষার দু'বার প্রয়োগের মাঝে সময়ের ব্যবধান থাকায় অভীক্ষার্থীদের পূর্ব অভিজ্ঞতা ও নতুন অর্জিত অভিজ্ঞতা অভীক্ষার ফলাফলকে প্রভাবিত করে, ফলে নির্ভরযোগ্যতার সঠিক পরিমাপ পাওয়া যায় না।

(3) **অভীক্ষক নির্ভর কারণ :**

অভীক্ষকের ব্যক্তিগত মতামত, জ্ঞান-অভিজ্ঞতা নির্ভরযোগ্যতা হ্রাসের অন্যতম কারণ। যেমন—

Scanned with CamScanner

(i) **ব্যক্তিগত মতামত (Personal Opinion) :** অভীক্ষাপত্র মূল্যায়নে অভীক্ষক তাঁর ব্যক্তিগত মতামত আরোপ করলে অভীক্ষার নির্ভরযোগ্যতা হ্রাস পেতে পারে।

(ii) **অভীক্ষা প্রয়োগ ও পরিচালনার অভিজ্ঞতা (Experience of Adminis-tering a Test) :** অভীক্ষা প্রয়োগ ও পরিচালনা সম্পর্কে অভীক্ষকের যথার্থ জ্ঞান ও অভিজ্ঞতা না থাকলে অভীক্ষার নির্ভরযোগ্যতা হ্রাস পেতে পারে।

অর্থাৎ, অভীক্ষা গঠন ও তার প্রয়োগের সময় উপরোক্ত বিষয়গুলিকে যথাযথভাবে নিয়ন্ত্রণ না করলে নির্ভরযোগ্যতার সঠিক পরিমাপ পাওয়া সম্ভব নয়।

২.৪ যথার্থতা : অর্থ, প্রকারভেদ ও পরিমাপ (Validity: Meaning, type and measurment) :

❑ ২.৪.১ সংজ্ঞা ও ধারণা (Definition and Concept) :

কোনো অভীক্ষা ব্যক্তির যে মানসিক বা শিক্ষাগত বৈশিষ্ট্যকে পরিমাপের জন্য প্রস্তুত করা হয়েছে, অভীক্ষাটি কতটা নিখুঁত ও সার্থকভাবে তা পরিমাপে সক্ষম হচ্ছে তার মাত্রাই হল অভীক্ষাটির **যথার্থতা (Validity)**।

অর্থাৎ, ভৌতবিজ্ঞানের একটি পারদর্শিতার অভীক্ষা যে মাত্রায় বা যতটা নির্ভুলভাবে শিক্ষার্থীর ভৌতবিজ্ঞানের জ্ঞান পরিমাপে সক্ষম, তাকেই বলা হবে অভীক্ষাটির **যথার্থতা** এবং যে সূচক দ্বারা এই মাত্রা প্রকাশ করা হয় তাকে **যথার্থতার সূচক (Index of Validity)** বলে।

Stanley and **Hopkins** (1978) এর মতে— "The validity of a measure is how well it fulfils the function for which it is being used– the degree to which it is capable of achieving certain aims."

N.E. Gronlund (1981) বলেছেন,— "Validity refers to the extent to which the results of an evaluation procedure serve the particular uses for which they are intended."

J.C. Aggarwal যথার্থতার সংজ্ঞায় বলেছেন,— "By validity we mean the degree to which a test or examination measures what it is intended to measure.

উপরের সংজ্ঞাগুলি বিশ্লেষণ করে বলা যায়, কোনো অভীক্ষা যতটা সার্থকভাবে ও নির্ভুলভাবে অভীষ্ট গুণকে পরিমাপ করতে পারে বা সক্ষম হয় তাই হল অভীক্ষাটির যথার্থতা।

Scanned with CamScanner